TECHNIQUES FOR ESTIMATING PEAK-FLOW FREQUENCY RELATIONS FOR NORTH DAKOTA STREAMS

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 92-4020

Prepared in cooperation with

NORTH DAKOTA DEPARTMENT OF TRANSPORTATION

TECHNIQUES FOR ESTIMATING PEAK-FLOW FREQUENCY RELATIONS FOR NORTH DAKOTA STREAMS

By Tara Williams-Sether

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 92-4020

Prepared in cooperation with

NORTH DAKOTA DEPARTMENT OF TRANSPORTATION

Bismarck, North Dakota

1992

U.S. DEPARTMENT OF THE INTERIOR

MANUEL LUJAN, JR., Secretary

U.S. GEOLOGICAL SURVEY

Dallas L. Peck, Director

For additional information
write to:

District Chief
U.S. Geological Survey
Water Resources Division
821 East Interstate Avenue
Bismarck, ND 58501

Copies of this report can be
purchased from:

U.S. Geological Survey
Books and Open-File Reports
Federal Center
Box 25425
Denver, CO 80225

CONTENTS

ILLUSTRATIONS

TABLES

CONVERSION FACTORS AND VERTICAL DATUM

Multiply	By	To obtain
cubic foot per second (ft^3/s)	0.02832	cubic meter per second
foot (ft)	0.3048	meter
foot per mile (ft/mi)	0.1894	meter per kilometer
inch (in.)	25.4	millimeter
inch per hour (in/hr)	25.4	millimeter per hour
mile (mi)	1.609	kilometer
square mile (mi^2)	2.590	square kilometer

To convert degrees Fahrenheit (°F) to degrees Celsius (°C), use the following formula: °C = 5/9 x (°F-32).

Sea level: In this report "sea level" refers to the National Geodetic Vertical Datum of 1929 (NGVD of 1929)--a geodetic datum derived from a general adjustment of the first-order level nets of both the United States and Canada, formerly called Sea Level Datum of 1929.

TECHNIQUES FOR ESTIMATING PEAK-FLOW FREQUENCY RELATIONS FOR NORTH DAKOTA STREAMS

By Tara Williams-Sether

ABSTRACT

This report presents techniques for estimating peak-flow frequency relations for North Dakota streams. In addition, a generalized skew coefficient analysis was completed for North Dakota to test the validity of using the generalized skew coefficient map in Bulletin 17B of the Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982, "Guidelines for Determining Flood Flow Frequency." The analysis indicates that the generalized skew coefficient map in Bulletin 17B provides accurate estimates of generalized skew coefficient values for natural-flow streams in North Dakota.

Peak-flow records through 1988 for 192 continuous- and partial-record streamflow gaging stations that had 10 or more years of record were used in a generalized least-squares regression analysis that relates peak flows for selected recurrence intervals to selected basin characteristics. Peak-flow equations were developed for recurrence intervals of 2, 10, 15, 25, 50, 100, and 500 years for three hydrologic regions in North Dakota. The peak-flow equations are applicable to natural-flow streams that have drainage areas of less than or equal to 1,000 square miles. The standard error of estimate for the three hydrologic regions ranges from 60 to 70 percent for the 100-year peak-flow equations.

Methods are presented for transferring peak-flow data from gaging stations to ungaged sites on the same stream and for determining peak flows for ungaged sites on ungaged streams. Peak-flow relations, weighted estimates of peak flow, and selected basin characteristics are tabulated for the 192 gaging stations used in the generalized skew coefficient and regression analyses. Peak-flow relations also are provided for 63 additional gaging stations that were not used in the generalized skew coefficient and regression analyses. These 63 gaging stations generally represent streams that are significantly controlled by regulation and those that have drainage areas greater than 1,000 square miles.

INTRODUCTION

The size of highway structures, such as bridges and culverts, is determined largely by the hydraulic properties necessary to convey peak streamflow without causing significant damage to upstream property or to the structures. Underdesign of a structure could cause disruption of service, costly maintenance, and loss of life. Overdesign of a structure could result in excessive construction costs. Thus, accurate estimates of peak streamflow for recurrence intervals such as the 15-, 25-, 50-, 100-, and 500-year peak flows are needed.

Estimates of peak flows for selected recurrence intervals for streamflow gaging stations commonly are determined by fitting a probability distribution function or a frequency relation to a series of annual peak flows. Estimates for ungaged sites commonly are made by transferring the data from gaging stations using regression techniques that relate peak flow for selected recurrence intervals to basin and climatic characteristics. Accuracy of the estimates for the ungaged sites depends on the accuracy of peak flows for the gaging stations for selected recurrence intervals, the accuracy of the basin and climatic characteristics, and the accuracy of the equations that relate peak flow to basin and climatic characteristics.

This report was prepared by the U.S. Geological Survey in cooperation with the North Dakota Department of Transportation. Engineering analyses performed by the North Dakota Department of Transportation require current peak-flow estimates, particularly for the 15- to 500-year recurrence intervals. These peak-flow requirements have created a need for updating the flood-frequency analysis by Crosby (1975). Crosby's analysis only used data through 1973 and did not include peak-flow equations for the 15-, 100-, and 500-year recurrence intervals. Because the accuracy of peak-flow estimates for selected recurrence intervals is sensitive to the value of the generalized skew coefficient, the generalized skew coefficient values presently being used in North Dakota also need to be evaluated.

All Federal agencies and many State agencies and local consultants use the log-Pearson Type III distribution when determining frequency relations for gaging stations and follow the procedures described in Bulletin 17B, "Guidelines for Determining Flood Flow Frequency" (Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982). In this report, this reference will be referred to as Bulletin 17B. The procedures described in Bulletin 17B were used to determine frequencies in this report.

Purpose and Scope

The purpose of this report is: (1) To test the validity of using the generalized skew coefficient map from Bulletin 17B for North Dakota and, if necessary, develop a new procedure for estimating the generalized skew coefficient for North Dakota streams; (2) to publish peak-flow frequency relations for selected recurrence intervals from 2 to 500 years for gaging stations with 10 or more years of record; (3) to present a procedure for computing peak-flow frequency relations for ungaged sites near a gaging station on a gaged stream; and (4) to develop regression equations for estimating peak flows for ungaged sites on ungaged natural-flow streams that have drainage areas of 1,000 mi^2 or less.

The generalized skew coefficient and regression analyses were based on data for 192 continuous- and partial-record streamflow gaging stations that had 10 or more years of record. Of these stations, 146 are in North Dakota, 14 are in Minnesota, 16 are in Montana, 13 are in South Dakota, 2 are in Manitoba, Canada, and 1 is in Saskatchewan, Canada. The locations of the 192 gaging stations are shown in figure 1. Log-Pearson Type III distribution was used to compute peak-flow frequency relations for the 192 gaging stations. The generalized least-squares (GLS) regression technique (Tasker and Stedinger, 1989) was used to compute weighted peak-flow frequency relations for the 192 stations and to develop equations that can be used to estimate peak-flow frequency relations for ungaged sites. A drainage-area ratio method was used to demonstrate how to estimate peak flow for ungaged sites near a gaging station on a gaged stream.

Previous Studies

McCabe and Crosby (1959) used all data available through 1955 and completed a study of the magnitude and frequency of floods in North Dakota and South Dakota. Patterson (1966) used data through 1961 and Patterson and Gamble (1968) used data through 1963 to complete part of a series of reports on the magnitude and frequency of floods in the United States. However, their reports contained little information on flood frequencies and magnitudes for small drainage basins in North Dakota. Crosby (1970) included a limited analysis of magnitude and frequency of floods when he evaluated the streamflow data program for North Dakota. Crosby (1975) used data available through 1973 to complete a study of the magnitude and frequency of floods for small drainage basins of 100 mi^2 or less; however, he did not develop equations for estimating the 15-, 100-, and 500-year recurrence interval floods. Miller and Frink (1984) completed a study to determine whether any changes in flood response because of changes in land use could be documented for the Red River of the North.

General Physical and Hydrologic Description of Study Area

The North Dakota landscape varies from flat plains to rugged badlands. Elevations range from 750 ft above sea level at Pembina in the extreme northeast corner of the State to 3,506 ft above sea level at White Butte near Bowman in the southwest corner. North Dakota commonly is divided into two major physiographic provinces, the Central Lowland and the Great Plains (Fenneman, 1946; fig. 2). The division between the provinces is along the easterly boundary of the Coteau du Missouri. The Central Lowland Province consists of the Red River Valley and the Drift Prairie. The Red River Valley is a flat area characterized by lake-bottom sediments near the Red River of the North and beach ridges and scarps farther west (Bluemle, 1977), and the Drift Prairie generally consists of a gently rolling area of glacial deposits. The Great Plains Province is a rolling to hilly area, the result of erosion of generally flat-lying, easily eroded sedimentary rocks. The Coteau du Missouri is a hummocky area over which glacial stagnation occurred. In the southwestern part of the State, badlands topography has developed near some

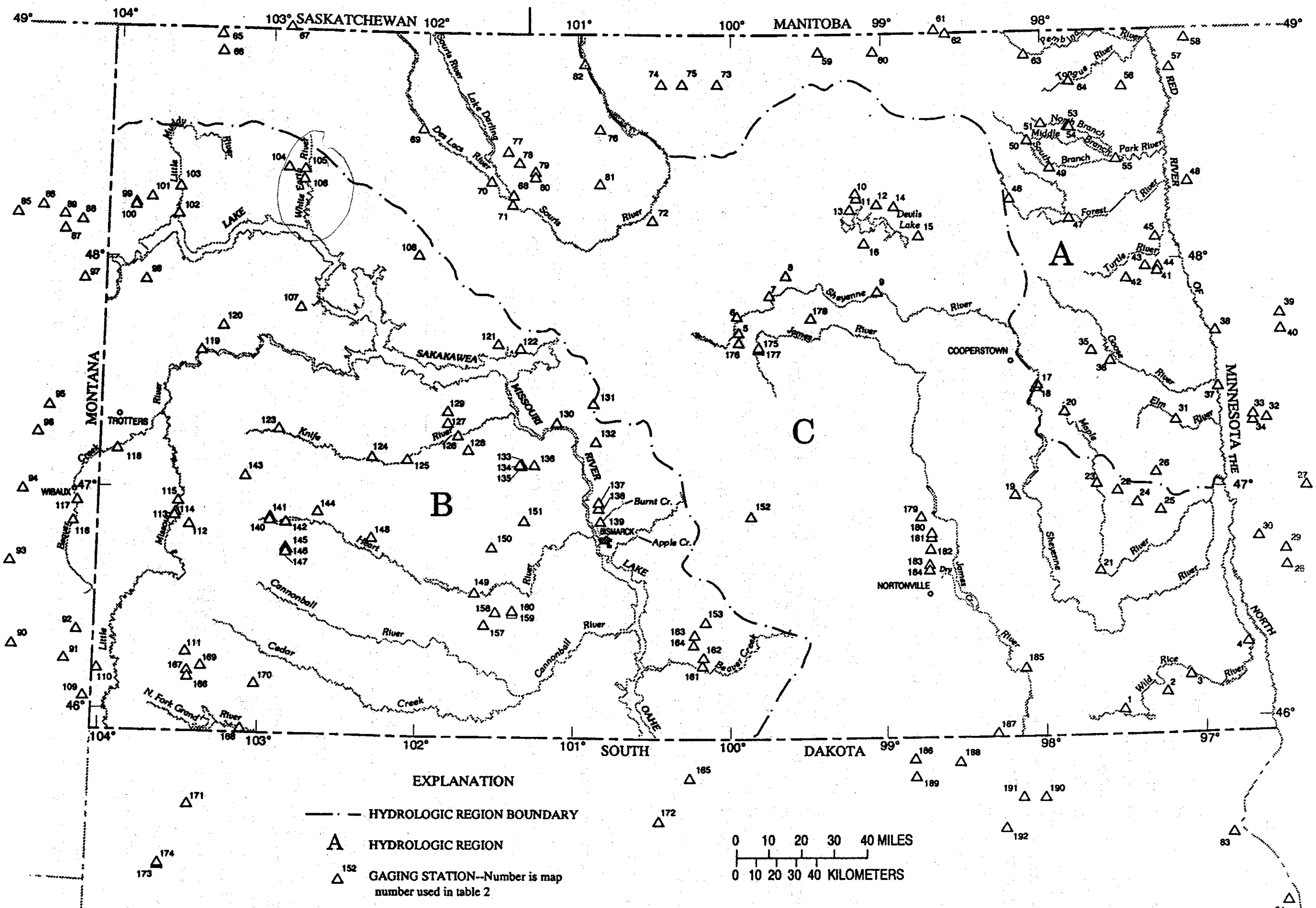

Figure 1.--Locations of gaging stations and regional hydrologic boundaries used in the regression analysis.

streams and rivers and is especially prominent along the Little Missouri River (Bluemle, 1977, p. 2). The two major drainage basins within North Dakota are the Red River of the North basin and the Missouri River basin (fig. 2).

North Dakota's climate has a great deal of variability compared to some states. Jensen (no date, p. 1) stated that "***because of location, the climate of the state is characterized by large annual, daily and day-to-day temperature changes, light to moderate precipitation which tends to be irregular in time and coverage, low relative humidity, plentiful sunshine, and nearly continuous air movement." Mean annual precipitation ranges from about 14 in. in the northwest part of the State to about 21.5 in. in the extreme southeast part of the State (Dr. J. W. Enz, North Dakota State Climatologist, written commun., 1988).

Peak flows in North Dakota result from both rainfall and snowmelt. About 75 percent of the precipitation occurs between April and September, and local high-intensity thunderstorms are common. Winter precipitation is minimal and generally in the form of snow (Jensen, no date). A comparison between the number of annual peak flows resulting from rainfall and the number of annual peak flows resulting from snowmelt for gaging stations located in the western part of the State indicated that peak flows resulting from snowmelt were more prevalent than those resulting from high-intensity rainfall (Emerson, 1988). A cursory comparison of annual peak flows made for gaging stations that were used in this peak-flow frequency analysis for the eastern part of the State indicated that the number of annual peak flows resulting from high-intensity rainfall and those resulting from snowmelt are about equal. No attempt was made to separate the peak flows resulting from rainfall from those resulting from snowmelt for this study because of the difficulty in distinguishing snowmelt peak flows that were not affected by rainfall. Also, such a separation could lead to an inadequate data base because 60 percent of the data base consists of record lengths of less than 20 years.

BASIN AND CLIMATIC CHARACTERISTICS CONSIDERED IN THE GENERALIZED SKEW COEFFICIENT AND REGRESSION ANALYSES

Basin and climatic characteristics were used in the analysis of generalized skew coefficients and in the development of regression equations for estimating flood characteristics for ungaged sites. Basin and climatic characteristics used in these analyses were obtained from the U.S. Geological Survey's National Water Data Storage and Retrieval System (WATSTORE) (Dempster, 1983). The characteristics were chosen on the basis of the results from previous studies and the availability of existing data. Methods for determining basin and climatic characteristics are discussed in more detail in Benson (1962) and Benson and Carter (1973). Climatological information for 1931-60 is available in Jensen (no date). The basin and climatic characteristics considered are:

1. Drainage area, *A*, in square miles--total area from gaging station to basin divide including noncontributing area.
2. Contributing drainage area, *CA*, in square miles--drainage area that contributes overland flow to the station in question.

3. Main-channel slope, *S*, in feet per mile--as computed by the elevation difference between points located 10- and 85-percent of the distance along the longest water course from the gaging station to the basin divide and divided by 0.75 of the longest water course length. Main-channel slope is further described by Benson (1962, p. B26-B28).

4. Stream length, *L*, in miles--measured along the longest water course from the gaging station to the basin divide including the main channel and appropriate upstream tributaries.

5. Mean basin elevation, *E*, in feet above sea level--measured from topographic maps by the transparent-grid sampling method; an average of 20 to 80 sample points in the basin.

6. Soil infiltration index, *SI*, in inches per hour--a relative measure of potential infiltration. Information on soil infiltration indices can be obtained from the Soil Conservation Service.

7. Mean annual precipitation, *P*, in inches.

8. Precipitation intensity, *I*, 2-year, 24-hour rainfall, in inches per 24 hours--estimated from the U.S. Weather Bureau (1961).

9. Mean annual snowfall, *SF*, in inches.

10. Mean maximum March air temperature, *M*, in degrees Fahrenheit.

ANALYSIS OF GENERALIZED SKEW COEFFICIENT

An analysis was undertaken to test the validity of using the generalized skew coefficient map from Bulletin 17B for North Dakota and to determine if an improved estimate of generalized skew could be obtained. The skew coefficient for the gaging-station record of annual peak flow is sensitive to extreme events; thus, it is difficult to obtain accurate gaging-station skew coefficient estimates from small samples of peak flow. The accuracy of the estimated gaging-station skew coefficient can be improved by weighting the gaging-station skew coefficient with a generalized skew coefficient estimated by considering information from nearby sites, as described in Bulletin 17B. The procedure for developing generalized skew coefficients requires the use of at least 40 gaging stations with 25 or more years of record for each station. The criteria of 25 or more years of record was relaxed to 10 or more years of record for this study in order to enlarge the available data base. Data for 192 gaging stations (fig. 1) were used in the analysis and the three methods described in Bulletin 17B were considered: (1) Lines of equal values of the skew coefficient drawn on a statewide map; (2) regression equations that relate skew coefficients to selected basin and climatic characteristics; and (3) the means of the gaging-station skew coefficent values for major river basins. Conditional probability adjustments (Appendix 5, Bulletin 17B) for low outliers and zero peak-flow records were applied to the station-record statistics. No adjustments were made for high outliers or historic peaks due to the paucity of historical information.

For method 1, each gaging-station skew coefficient value was plotted on a map of North Dakota at the centroid of the drainage basin, and the map was examined to determine whether any geographic or topographic trends were apparent. Very little relation to geography or topography was apparent, so this method was not considered further.

For method 2, plots of gaging-station skew coefficient values versus selected basin and climatic characteristics were generated. Development of regression equations was attempted using all of the skew coefficient data for the 192 gaging stations, using only data for gaging stations located in the Red River of the North basin, and using only data for gaging stations located in the Missouri River basin. The regression equations that relate gaging-station skew coefficient values to basin and climatic characteristics were not well defined and, consequently, this method was not considered further.

For method 3, gaging-station skew coefficient values were derived for each gaging station for three categories; stations with 10 or more years of record, stations with 10 to 19 years of record, and stations with 20 or more years of record. A mean skew coefficient value, the generalized skew coefficient, was computed for each category using all of the 192 gaging stations, those in the Red River of the North basin, and those in the Missouri River basin. The mean-squared-error (MSE) value associated with the nine computed generalized skew coefficients was compared with the MSE value of 0.302 for the generalized skew coefficient map in Bulletin 17B. Improvement in MSE was most noticeable for gaging stations that have 20 or more years of data, particularly for the data set that represents only the Red River of the North basin and for the data set that represents all of the 192 stations (table 1). However, the change in MSE was not considered to be significant enough to change the common procedure of using the generalized skew coefficient map in Bulletin 17B for North Dakota streams. Thus, it was determined that the generalized skew coefficient map in Bulletin 17B provides accurate estimates of generalized skew coefficient values for natural-flow streams in North Dakota. The generalized skew coefficient map mean of -0.4 (Bulletin 17B) is considered representative for North Dakota as an average. Values obtained from the generalized skew coefficient map in Bulletin 17B are considered more accurate for individual stations, especially in the southwestern corner of the State.

PEAK-FLOW FREQUENCY RELATIONS FOR GAGING STATIONS

Peak-flow frequency relations for 192 gaging stations on streams that have drainage areas of less than or equal to 1,000 mi^2 are listed in table 2 (at the back of report). The first-line values listed in table 2 are the gaging-station peak-flow frequency data. These frequency data were determined using the procedures outlined in Bulletin 17B. The second line values are the weighted peak-flow frequency values that were determined using the gaging-station peak-flow frequency data and an estimate of peak-flow frequency from regional regression equations (discussed later). The weighted value is considered to be the best estimate of the peak-flow frequency relation for a gaging station on an unregulated stream because the weighted estimate reduces the time-sampling error associated with short periods of record.

Peak-flow frequency relations for 63 additional gaging stations that were not used in analysis of the generalized skew coefficient or for developing regression equations for computing peak-flow frequencies at ungaged sites are listed in table 3 (at the back of report). These stations generally represent streams that are significantly controlled by regulation and those with

Table 1.--Mean and variance of station skew data and mean squared error computed for the differences between station skew and the generalized skew coefficient map in Bulletin 17B

Record length for stations (years)	Number of stations	Mean of station skew	Variance of station skew	Mean-squared error of the difference between station skew and Bulletin 17B skew
		All 192 stations		
10 or more	192	-0.378	0.358	0.353
10 to 19	108	-.324	.393	.388
20 or more	84	-.483	.296	.301
		Missouri River basin stations only		
10 or more	108	-0.269	0.382	0.387
10 to 19	72	-.191	.385	.411
20 or more	36	-.425	.349	.338
		Red River of the North basin stations only		
10 or more	84	-0.517	0.297	0.310
10 to 19	36	-.536	.394	.402
20 or more	48	-.515	.231	.244

drainage areas greater than 1,000 mi^2. Locations of these gaging stations are shown in figure 3 (at the back of report).

It is good practice to coordinate any work related to peak-flow frequency with other government agencies and appropriate consulting firms. This is especially important for sites along state borders and on the larger rivers for which other agencies and consulting firms may have determined peak-flow values while completing flood-plain information or flood-insurance or flood-frequency studies. Also, coordinating the peak-flow frequency values with the respective agencies is needed to determine whether the operating rules for regulating the streams have been consistent throughout the period of record used to determine the peak-flow frequency values listed in table 3. The North Dakota State Water Commission, U.S. Army Corps of Engineers, U.S. Geological Survey, U.S. Soil Conservation Service, U.S. Bureau of Reclamation, and U.S. Bureau of Land Management are agencies that commonly develop peak-flow information for selected parts of the State. The peak-flow values listed in table 3 were coordinated with the North Dakota State Water Commission and the U.S. Army Corps of Engineers (K. L. Lindskov, U.S. Geological Survey, written commun., 1991).

The weighting procedure used to determine weighted peak flow in this report is part of the GLS regression and network analysis program GLSNET (G. D. Tasker and others, written commun., 1987). The following equation was used to calculate the weighted peak flows for each gaging station in table 2:

$$Q_{TW} = \frac{n\,Q_{TS} + en\,Q_{TR}}{n + en}, \tag{1}$$

where

Q_{TW} = weighted peak flow, in cubic feet per second, for recurrence interval of T-years;

n = number of years of station data used to compute Q_{TS};

Q_{TS} = station peak flow, in cubic feet per second, for recurrence interval of T-years;

en = equivalent years of record for Q_{TR}; and

Q_{TR} = peak flow from regional regression equation that relates peak flow to basin characteristics, in cubic feet per second, for recurrence interval of T-years.

The equivalent years of record, en, is a measure of the accuracy of prediction in terms of the number of years of record that is required for each gaging station to achieve results of equal accuracy to that of the regional regression equation. A further explanation on how the equivalent years of record is calculated is given by Hardison (1971).

PEAK-FLOW FREQUENCY RELATIONS FOR UNGAGED SITES

The procedure for determining peak flow for a selected recurrence interval for a specific ungaged site depends on whether the site is located near a gaging station on the same stream or is an ungaged site on an ungaged stream. For an ungaged site near a gaging station on the same stream, a drainage-area ratio method should be used to estimate the T-year peak flow. For ungaged sites on ungaged streams, the regional regression equations developed during this study that relate peak flow for selected recurrence intervals to basin characteristics should be used.

Ungaged Sites Near a Gaging Station on the Same Stream

The following equation can be used to determine T-year peak-flow values for an ungaged site located near a gaging station on the same stream. This equation should be used only when the contributing drainage area for the ungaged site is from 75 to 150 percent of the contributing drainage area for the gaged site (otherwise the regional regression equations discussed later in this report should be used):

$$Q_{T(u)} = Q_{TW(g)}\ (CA_u/CA_g)^x, \quad (2)$$

where

$Q_{T(u)}$ = peak flow, in cubic feet per second, for the ungaged site for a recurrence interval of T-years;

$Q_{TW(g)}$ = weighted peak flow, in cubic feet per second, for the gaging station for a recurrence interval of T-years;

CA_u = contributing drainage area, in square miles, for the ungaged site;

CA_g = contributing drainage area, in square miles, for the gaging station; and

x = mean exponent for the appropriate hydrologic region.

The peak flow, $Q_{TW(g)}$, is the weighted value in table 2 that was determined using equation 1. The exponent, x, was determined by relating the seven separate regression equations for T-year peak flows (T = 2, 10, 15, 25, 50, 100, 500 years) to the independent drainage-area variable for each hydrologic region and computing the mean of the exponents for the seven equations for each region. The mean exponent for region A was 0.55, the mean exponent for region B was 0.58, and the mean exponent for region C was 0.57. The determination of the hydrologic regions and development of the regional regression equations are discussed in the section "Ungaged Sites on Ungaged Streams."

Ungaged Sites on Ungaged Streams

Development of Regression Equations

Equations that relate peak-flow values for selected recurrence intervals to basin characteristics were developed using a GLS regression procedure instead of using the conventional ordinary least-squares (OLS) regression procedure. The GLS regression procedure takes into consideration the time-sampling error in the dependent variable and the cross correlation between sites. When available peak-flow data for gaging stations are for different and widely varying lengths of record and concurrent peak flows at different sites are cross correlated, the GLS regression procedure provides more accurate estimates of the regression coefficients. The GLS regression procedure also provides better estimates of the predictive accuracy of peak-flow values that are computed by the regression equations and almost unbiased estimates of the variance of the underlying regression model error (Stedinger and Tasker, 1985).

The conventional OLS regression procedure, such as that used by Crosby (1975), only provides an estimate of the model error, and the assumption is made that peak flows for selected recurrence intervals are independent from site to site and that site-to-site variance (time-sampling error) is identical. The OLS procedure, however, was used in this study to tentatively choose the basin and climatic characteristics that would be statistically significant at the 5-percent level. Both dependent and independent variables were transformed to base 10 logarithms for the regression analyses. The transformation to base 10 logarithms was made to linearize the relation between the dependent and one or more independent variables.

The basic regression model using the GLS procedure can be represented by the following linear equation:

$$\underline{Y} = \underline{x}\ \underline{B} + \underline{u}, \tag{3}$$

where

$\underline{Y}$ = ($nx1$) vector of T-year flood events (dependent variable),
$\underline{x}$ = (nxp) matrix of basin characteristics (independent variables),
$\underline{B}$ = ($px1$) vector to be estimated (regression coefficients), and
$\underline{u}$ = ($nx1$) random vector (errors).

The best linear unbiased estimator ($\underline{b}$) of the parameter vector $\underline{B}$ for the T-year event is:

$$\underline{b} = (\underline{x}^T\ \underline{\Lambda}^{-1}\ \underline{x})^{-1}\ \underline{x}^T\ \underline{\Lambda}^{-1}\ \underline{Y},$$

where

$\underline{\Lambda}$ = the unknown covariance (weighting) matrix.

Further details on the GLS procedure and the methods used to determine the unknown matrix are discussed in Stedinger and Tasker (1985) and in Tasker and Stedinger (1989). The computer program GLSNET (G. D. Tasker and others, written commun., 1987) was used to develop the T-year peak-flow regional regression equations for North Dakota.

An evaluation of residuals from preliminary regression runs showed that the State should be divided into the same three hydrologic regions (fig. 1) that were used in the previous study by Crosby (1975). The boundaries of the hydrologic regions generally conform to three primary types of topography: Region A is a flat plain area, region B is a rolling to hilly plain area, and region C is a prairie pothole area.

Regression equations for recurrence intervals of 2, 10, 15, 25, 50, 100, and 500 years were developed for hydrologic regions A, B, and C. The characteristics used in the final regional regression equations that are significant at the 5-percent level are contributing drainage area and main-channel slope. The regional regression equations and the associated standard error of estimate, standard error of prediction, and equivalent years of record for each equation are listed in table 4. The standard error of estimate is the square root of the average model error variance in the GLS analysis. The model error variance is a measure of the error inherent in the regression model that cannot be changed by collecting more data. The standard error of prediction is the square root of the sum of the average sampling error variance and the average model error variance and is the measure of the accuracy with which the regression model can estimate the T-year flood at an ungaged site. The equivalent years of record indicates the number of years of streamflow record that provides an estimate equal in accuracy to the standard error of prediction. The regression equations developed for regions A, B, and C may be used to determine peak flows for ungaged sites on ungaged streams. Examples of the use of the regional regression equations to compute peak-flow values for ungaged sites on ungaged streams are given in the section "Examples of Using Regression Equations to Compute Peak-Flow Frequency Data for Ungaged Sites on Ungaged Streams."

Table 4.--Regional regression equations that relate peak flow for selected recurrence intervals to selected basin characteristics

[Q, peak flow, in cubic feet per second; CA, contributing drainage area, in square miles; S, main-channel slope, in feet per mile]

Recurrence interval (years)	Equation	Number of stations used in analysis	Standard error of estimate (percent)	Standard error of prediction (percent)	Equivalent years of record (years)
		Region A			
2	$Q = 24.9\ CA^{0.543}\ S^{0.094}$	41	60	64	3.1
10	$Q = 62.2\ CA^{0.600}\ S^{0.168}$	41	55	60	5.0
15	$Q = 70.9\ CA^{0.609}\ S^{0.181}$	41	56	60	5.6
25	$Q = 81.6\ CA^{0.619}\ S^{0.197}$	41	57	61	6.3
50	$Q = 95.9\ CA^{0.631}\ S^{0.217}$	41	58	64	7.1
100	$Q = 110\ CA^{0.640}\ S^{0.234}$	41	60	66	7.8
500	$Q = 142\ CA^{0.656}\ S^{0.268}$	41	67	73	8.7
		Region B			
2	$Q = 7.68\ CA^{0.697}\ S^{0.299}$	88	83	88	2.3
10	$Q = 32.7\ CA^{0.716}\ S^{0.294}$	88	60	64	5.9
15	$Q = 41.6\ CA^{0.717}\ S^{0.286}$	88	60	67	6.7
25	$Q = 55.1\ CA^{0.716}\ S^{0.276}$	88	61	66	7.5
50	$Q = 76.4\ CA^{0.715}\ S^{0.262}$	88	65	70	8.2
100	$Q = 101\ CA^{0.713}\ S^{0.249}$	88	70	76	8.5
500	$Q = 171\ CA^{0.708}\ S^{0.229}$	88	84	91	8.6
		Region C			
2	$Q = 4.08\ CA^{0.638}\ S^{0.348}$	58	98	104	2.0
10	$Q = 22.3\ CA^{0.665}\ S^{0.275}$	58	66	71	5.2
15	$Q = 29.4\ CA^{0.668}\ S^{0.263}$	58	64	77	6.3
25	$Q = 39.7\ CA^{0.670}\ S^{0.249}$	58	62	68	7.5
50	$Q = 56.3\ CA^{0.671}\ S^{0.232}$	58	62	68	9.0
100	$Q = 75.6\ CA^{0.672}\ S^{0.219}$	58	63	69	10.2
500	$Q = 129\ CA^{0.676}\ S^{0.196}$	58	67	75	12.0

Limitations on Use of the Regression Equations

The following limitations should be considered when using the regression equations to compute peak-flow frequency relations for North Dakota streams: (1) The equations apply to streams that are located in rural watersheds and should not be applied to watersheds significantly affected by urbanization; (2) the equations should not be used where dams, flood-detention structures, and other manmade works exist that significantly affect the annual peak flows; and (3) the equations generally should be used only for streams that have drainage areas of less than or equal to 1,000 mi^2 and for streams that have drainage areas and basin characteristics that are within the range of characteristics used to develop the regression equations. The ranges of the characteristics for the gaging stations used to develop the regression equations are given in table 5.

Table 5.--Range of basin characteristics used to develop the regression equations

Region	Contributing drainage area (square miles) (*CA*)	Main-channel slope (feet per mile) (*S*)
A	0.08 - 959	1.0 - 300
B	0.06 - 775	2.1 - 198
C	0.13 - 796	2.1 - 270

EXAMPLES OF ESTIMATING PEAK-FLOW FREQUENCY DATA FOR GAGING STATIONS AND FOR UNGAGED SITES

Example for Determining Weighted Peak-Flow Frequency Data for Gaging Stations

The procedure used to determine the weighted peak flow (eq. 1) for the 100-year recurrence interval for station 05056950, Sheyenne River tributary no. 2 near Cooperstown, N.Dak. (table 2, map number 18), is shown in the following example.

Station 05056950,
n = 15 years,
Q_{TS} = 98 ft^3/s,
CA = 0.08 mi^2, and
S = 300 ft/mi.

The gaging station is located in region A, and the 100-year regression equation from table 4 is

$$Q_{100} = 110\ CA^{0.640}\ S^{0.234};\ \text{thus,}$$

$$Q_{TR} = Q_{100} = 110\ (0.08)^{0.640}\ (300)^{0.234} = 83.$$

The equivalent years of record, *en*, for this station for the 100-year equation is 3.8 years. Using equation 1,

$$Q_{TW} = \frac{15(98) + 3.8(83)}{15 + 3.8} = 95\ ft^3/s.$$

Example for Determining Peak-Flow Frequency Data for an Ungaged Site Near a Gaging Station on the Same Stream

The following is an example of how to use the drainage-area ratio method (eq. 2) to determine peak flow for an ungaged site near a gaging station on the same stream. To estimate the 100-year peak flow for an ungaged site on Burnt Creek, locate the site in figure 1. The site is located downstream of a gaging station on the same stream (station 06342450, Burnt Creek near Bismarck, N.Dak.). Contributing drainage area (CA_u) for the ungaged site is 114 mi^2, and the contributing drainage area (CA_g) for the gaging station (table 2, map number 139) is 108 mi^2. Both the ungaged site and the gaging station are located in region B. Determine if the drainage-area ratio (CA_u/CA_g) is between 0.75 and 1.50:

$$CA_u/CA_g = 114/108 = 1.06,$$

which meets the drainage-area ratio requirement. Thus, the following relation is used:

$$Q_{100(u)} = Q_{100(g)}\ (CA_u/CA_g)^{0.58},$$

where
$Q_{100(g)}$ = 7,790 ft^3/s, the weighted peak flow for the gaging station (table 2);
CA_u = 114 mi^2; and
CA_g = 108 mi^2.
Therefore,

$$Q_{100(u)} = 7{,}790\ (114/108)^{0.58} = 8{,}040\ ft^3/s.$$

Examples for Determining Peak-Flow Frequency Data for an Ungaged Site

Between Two Gaging Stations on the Same Stream

An ungaged site for which a peak-flow calculation is desired sometimes may be between two gaging stations on the same stream. If the contributing drainage area for the ungaged site is within 75 to 150 percent of the contributing drainage area for both of the gaging stations, the drainage-area ratio method (eq. 2) should be applied to determine peak flow ($Q_{T(u)}$) for the ungaged site using data from each of the gaging stations. The resulting two peak-flow values then are averaged in log units. The peak flow ($Q_{T(u)}$) also can be determined by interpolating in log units between both of the gaging stations.

The following examples illustrate how to use equation 2 and interpolation to determine a 100-year peak flow for an ungaged site that is between two gaging stations on the same stream. The gaging stations are station 06336500, Beaver Creek at Wibaux, Mont. (table 2, map number 117), and station 06336600, Beaver Creek near Trotters, N.Dak. (table 2, map number 118). Contributing drainage area (CA_u) for the ungaged site is 500 mi^2, contributing drainage area (CA_g) for gaging station 06336500 is 351 mi^2, and contributing drainage area (CA_g) for gaging station 06336600 is 616 mi^2. The ungaged site and the gaging stations are located in region B.

Determine if the drainage-area ratio (CA_u/CA_g) for the ungaged site and each gaging station is between 0.75 and 1.50:

1. Station 06336500,
   CA_g = 351 mi^2,
   CA_u = 500 mi^2,
   CA_u/CA_g = 500/351 = 1.42.
2. Station 06336600,
   CA_g = 616 mi^2,
   CA_u = 500 mi^2,
   CA_u/CA_g = 500/616 = 0.81.

The drainage-area ratio requirement has been met for both situations. Thus, equation 2 is applied for the ungaged site using data from both gaging stations, and the results are averaged in log units.

1. Station 06336500,
   $Q_{100(g)}$ = 10,200 ft^3/s, weighted peak flow for the gaging station (table 2);
   $Q_{100(u)} = 10{,}200\ (500/351)^{0.58} = 12{,}500\ ft^3/s$.
2. Station 06336600,
   $Q_{100(g)}$ = 11,700 ft^3, weighted peak flow for the gaging station (table 2);
   $Q_{100(u)} = 11{,}700\ (500/616)^{0.58} = 10{,}400\ ft^3/s$.

Logarithmic average:

$$\frac{\log(12{,}500) + \log(10{,}400)}{2} = \frac{4.097 + 4.017}{2} = 4.057,$$

$$Q_{100(u)} = \text{antilog}(4.057) = 11{,}400 \text{ ft}^3/\text{s}.$$

The peak flow for the ungaged site determined by interpolation between the two gaging stations is shown below:

$$\log Q_{100(u)} = \log Q_{100(06336600)} - \frac{\log CA_{06336600} - \log CA_u}{\log CA_{06336600} - \log CA_{06336500}} \times [\log Q_{100(06336600)} - \log Q_{100(06336500)}]$$

$$\log Q_{100(u)} = \log(11{,}700) - \frac{\log(616) - \log(500)}{\log(616) - \log(351)} [\log(11{,}700) - \log(10{,}200)]$$

$$\log Q_{100(u)} = 4.068 - \frac{2.789 - 2.699}{2.789 - 2.545} (4.068 - 4.009) = 4.046$$

$$Q_{100(u)} = \text{antilog}(4.046) = 11{,}100 \text{ ft}^3/\text{s}.$$

Examples of Using Regression Equations to Compute Peak-Flow Frequency Data for Ungaged Sites on Ungaged Streams

Depending on the location of the ungaged site and its relation to the hydrologic region boundaries, one of two procedures can be used to compute peak-flow values. Procedure 1 should be used when the stream drainage area is within a single region. Procedure 2 should be used when the stream drainage area is part of more than one region.

Use procedure 1 to determine the 100-year peak flow for an ungaged site on Dry Creek near Nortonville, N.Dak. The site is located in region C (fig. 1). The basin characteristics that were computed from a U.S. Geological Survey 7.5-minute topographic map are:

CA = 8.0 mi^2 and
S = 4.2 ft/mi.

The indicated peak flow for a recurrence interval of 100 years is

$$Q_{100} = 75.6\ (8.0)^{0.672}\ (4.2)^{0.219} = 419 \text{ ft}^3/\text{s}.$$

Procedure 2 is similar to procedure 1 except T-year regional regression equations would be solved for each of the associated regions and the results would be averaged or apportioned according to the fraction of the contributing drainage area that is in each region. Because the boundaries for the three hydrologic regions A, B, and C were delineated generally along drainage divides of major streams, few streams cross these regional boundaries.

NEED FOR ADDITIONAL STREAMFLOW DATA AND FOR UPDATING THE BASIN AND CLIMATIC CHARACTERISTICS FILE

Future peak-flow frequency studies for North Dakota could be improved if additional gaging stations were established or previously discontinued gaging stations were reestablished on natural-flow streams, especially for sites on streams with drainage areas of less than 200 mi^2. Most of the gaging stations operated in 1990 are on larger streams and are not operated primarily for defining natural variations in hydrologic characteristics. In addition, the basin and climatic characteristics file needs to be updated.

Descriptions of the history of the streamflow-data program for North Dakota appear in Crosby (1970) and Ryan (1985, 1989). During 1954-73, as many as 92 crest-stage gages were operated to determine peak-flow frequency for sites on streams with small drainage areas (Crosby, 1975; Crosby and Pewe, 1972); and, at times during 1977-83, about 50 continuous-record gaging stations were operated to define hydrologic characteristics in the coal areas of North Dakota (Haffield, 1981). Currently (1990), 94 gaging stations for which annual peak flow is determined are operated in North Dakota; however, many of these 94 stations are on regulated streams. Additional stations need to be operated for defining natural variations in hydrologic characteristics, such as for determining peak-flow frequency at ungaged sites. The peak-flow information could be obtained by operating continuous-record gaging stations and (or) crest-stage gages for at least 25 years, an adequate record length for estimating the 100-year peak flow. The GLS procedure could be used to perform a network analysis that would yield a relation between the effect of adding or deleting stations to an operating budget. The GLS procedure also would relate the standard error (at site or regional) to record length with the idea of achieving a certain standard error (at site or regional).

Accurate and up-to-date data for basin and climatic characteristics also are an important part of defining relations for transferring peak-flow values from gaged to ungaged sites. Many U.S. Geological Survey 7.5-minute topographic maps have been completed since the basin and climatic characteristics were calculated for the study by Crosby (1970). Thus, the basin and climatic characteristics file for North Dakota needs to be updated and, in many cases, additional basin and climatic characteristics (beyond contributing drainage area and main-channel slope) need to be determined for gaging stations that were established after 1970.

SUMMARY

The generalized skew coefficient for North Dakota in Bulletin 17B of the Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982, "Guidelines for Determining Flood Flow Frequency" was evaluated for applicability by using the three recommended procedures: (1) Lines of equal skew coefficient drawn on a statewide map, (2) regression equations that relate skew coefficient to selected basin and climatic characteristics, and (3) mean values of skew coefficient for major river basins or hydrologic regions. Data for 192 continuous- and partial-record streamflow gaging stations that had 10 or more years of record were used to evaluate the use of the generalized skew coefficient map in Bulletin 17B and to compute peak-flow frequency relations. Of these gaging stations, 146 are in North Dakota, 14 in Minnesota, 16 in Montana, 13 in South Dakota, 2 in Manitoba, Canada, and 1 in Saskatchewan, Canada. The gaging stations were restricted to those on natural-flow streams not significantly affected by regulation that have drainage areas of less than or equal to 1,000 square miles. Interpretation of the data indicates that the generalized skew coefficient map for North Dakota in Bulletin 17B could not be improved significantly. Thus, it was determined that the generalized skew map provides accurate estimates of generalized skew coefficient values for natural-flow streams in North Dakota.

Peak-flow frequency relations for the 192 gaging stations used in the regression analysis were updated using data through 1988. The frequency relations are tabulated along with selected basin characteristics. The selected basin characteristics used to develop the peak-flow frequency relations and the generalized least-squares regression equations are contributing drainage area and main-channel slope. Weighted peak flows based on gaging-station data and the regional regression equations for selected recurrence intervals also are provided for each gaging station. The weighted value is considered to be the most reliable estimate of peak flow for a gaging station on an unregulated stream because of the reduction in time-sampling error associated with short periods of record. Peak-flow relations are provided for 63 additional gaging stations. These 63 stations generally represent streams that are significantly controlled by regulation and those with drainage areas greater than 1,000 square miles.

The generalized least-squares regression method was used to develop regression equations for three hydrologic regions in North Dakota by relating peak flows for recurrence intervals of 2, 10, 15, 25, 50, 100, and 500 years to basin characteristics. The standard error of estimate ranges from 60 to 70 percent for the 100-year peak-flow equations.

Examples are given for (1) determining weighted peak-flow frequency data for gaging stations and (2) using the drainage-area ratio method for determining peak-flow frequency data for ungaged sites near a gaging station on the same stream and ungaged sites between two gaging stations on the same stream. Also, two procedures are given for determining peak-flow values for ungaged sites on ungaged streams using the regional regression equations.

Future peak-flow frequency studies for North Dakota could be improved if additional gaging stations were established on natural-flow streams. Additional peak-flow frequency data are needed, especially for sites on streams with drainage areas of less than 200 square miles. The basin characteristics file needs to be updated and additional basin and climatic characteristics need to be determined for gaging stations that were established after 1970.

REFERENCES CITED

Benson, M.A., 1962, Factors influencing the occurrence of floods in a humid region of diverse terrain: U.S. Geological Survey Water-Supply Paper 1580-B, 64 p.

Benson, M.A., and Carter, R.W., 1973, A national study of the streamflow data-collection program: U.S. Geological Survey Water-Supply Paper 2028, 44 p.

Bluemle, J.P., 1977, The face of North Dakota, The geologic story: North Dakota Geological Survey Educational Series 11, 73 p.

Crosby, O.A., 1970, A proposed streamflow data program for North Dakota: U.S. Geological Survey open-file report, 68 p.

-----1975, Magnitude and frequency of floods in small drainage basins in North Dakota: U.S. Geological Survey Water-Resources Investigations 19-75, 24 p.

Crosby, O.A., and Pewe, R.A., 1972, Flood stages and discharges for small streams in North Dakota--Progress report to the North Dakota State Highway Department for the period July 1, 1954 to September 30, 1971: U.S. Geological Survey open-file report, 160 p.

Dempster, G.R., 1983, Streamflow basin characteristics, Chapter II, Section A, in WATSTORE user's guide, v. 4: U.S. Geological Survey Open-File Report 79-1336-I, p. A-24 to A-34.

Emerson, D.G., 1988, Surface-water hydrology of Hay Creek watershed, Montana, and West Branch Antelope Creek watershed, North Dakota: U.S. Geological Survey Water-Resources Investigations Report 88-4038, 111 p.

Fenneman, N.M., 1946, Physical divisions of the United States: U.S. Geological Survey map prepared in cooperation with the Physiographic Commission, U.S. Geological Survey, scale 1:700,000 (reprinted, 1964).

Haffield, N.D., 1981, Statistical summaries of streamflow and water-quality data for streams of western North Dakota, 1977-80: U.S. Geological Survey Open-File Report 81-1066, 78 p.

Hardison, C.H., 1971, Prediction error of regression estimates of streamflow characteristics at ungaged streams: U.S. Geological Survey Professional Paper 750-C, p. C228-C236.

Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood flow frequency: Hydrology Subcommittee Bulletin 17B, 28 p., 14 appendices.

Jensen, R.E., [no date], Climate of North Dakota: Fargo, North Dakota State University, 48 p.

McCabe, J.A., and Crosby, O.A., 1959, Floods in North and South Dakota, Frequency and magnitude: U.S. Geological Survey open-file report, 132 p.

Miller, J.E., and Frink, D.L., 1984, Changes in flood response of the Red River of the North basin, North Dakota-Minnesota: U.S. Geological Survey Water-Supply Paper 2243, 103 p.

Patterson, J.L., 1966, Magnitude and frequency of floods in the United States, pt. 6-A, Missouri River basin above Sioux City, Iowa: U.S. Geological Survey Water-Supply Paper 1679, 471 p.

Patterson, J.L., and Gamble, C.R., 1968, Magnitude and frequency of floods in the United States, pt. 5, Hudson Bay and upper Mississippi River basins: U.S. Geological Survey Water-Supply Paper 1678, 546 p.

Ryan, G.L., 1985, Data uses and funding of the streamflow-gaging program in North Dakota: U.S. Geological Survey Open-File Report 85-349, 29 p.

-----1989, Cost effectiveness of the stream-gaging program in North Dakota: U.S. Geological Survey Water-Resources Investigations Report 89-4068, 34 p.

Stedinger, J.R., and Tasker, G.D., 1985, Regional hydrologic analysis 1: Ordinary, weighted, and generalized least squares compared: Water Resources Research, v. 21, no. 9, p. 1421-1432.

Tasker, G.D., and Stedinger, J.R., 1989, An operational GLS model for hydrologic regression: Journal of Hydrology, v. 111, p. 361-375.

U.S. Weather Bureau, 1961, Rainfall frequency atlas of the United States: Technical Paper 40, 115 p.